

One hundred years of Current Employment Statistics data collection

To help mark the Monthly Labor Review's centennial, the editors invited several producers and users of BLS data to take a look back at the last 100 years. This article describes the evolution of data collection methods used by the Current Employment Statistics (CES) program. Over the past century, the program has made major strides in developing new survey instruments and processes, all designed to improve data quality and minimize respondent burden.

Current Employment Statistics (CES) survey data,¹ our nation's primary read on the economic heartbeat of employers, are released by the Bureau of Labor Statistics (BLS) at the beginning of each month. The data are part of the Employment Situation, commonly referred to as the Jobs Report.² This Principal Federal Economic Indicator celebrated its 100th anniversary in October 2015. Although the CES survey started in 1915 as a modest establishment survey of employment, earnings, and hours for four U.S. manufacturing industries, by 2015 it had grown to cover all nonfarm industries.³

CES data collection has evolved over the past century to incorporate a number of collection instruments, all tailored to efficiently gather monthly employment data on a large scale. The CES survey—a multimodal survey believed to be the largest of its type in the world—is designed to maximize survey response and data quality while minimizing cost and respondent burden. As a quick-response survey,⁴ it requires substantial resources to collect high-quality data before first preliminary estimates are calculated. By definition, the CES program collects information on employment, earnings, and hours for payrolls that include the 12th of the month. Preliminary national estimates are then released at the beginning of the subsequent month (typically on the first Friday of the month).⁵ CES has between 10 and 16 business days to collect data before the cutoff for calculation of initial estimates. Collection begins on the first business day after the 12th of the month and continues until the Monday of release week. Data collected after the closing date for initial estimates can be included in subsequent revisions.



Nicholas Johnson

johnson.nicholas@bls.gov

Nicholas Johnson is a supervisory economist in the Office of Employment and Unemployment Statistics, U.S. Bureau of Labor Statistics.

Monthly data collection primarily occurs during the second half of the month. Data collection centers (DCCs) in Atlanta, Dallas, Kansas City, and Fort Walton Beach, FL, fill with a flurry of activity, as BLS data collection agents contact 85,000 worksites nationwide. Concurrently, analysts at the BLS regional office in Chicago are parsing payroll data files provided electronically by some of the country's largest employers, representing nearly 130,000 worksites. Secure servers in Washington, DC, come to life as nearly 45,000 worksites report their employment data through Web-based applications and another 7,000 by touchtone data entry (TDE), an automated phone process. Further, state Labor Market Information agencies collect an additional 7,000 reports for CES, the vast majority of which are state government worksites.

As a celebration of the 100th anniversary of CES, this article reviews the development of the program's data collection procedures and instruments over the past century. The discussion tracks the transition from single-mode to multimode CES data collection.

Collecting CES data by mail (1915–83)

The CES program was established in October 1915 to collect, calculate, and publish statistics on monthly employment, hours, and earnings. Originally, the program surveyed only four industries. By the end of 1916, however, it had expanded its coverage to include 13 industries, representing a sample of 574 establishments.

Figure 1. CES data collection form, 1944

(B. L. S. 780) Budget Bureau No. 44-1286.
Approval expires January 15, 1945.

Manufacturing **CONFIDENTIAL**

THIS IS OUR PERMANENT OFFICE RECORD—PLEASE HANDLE CAREFULLY

B.L.S. Codes	No. Est.	State	Ind.	Est. No.	Ind. Grp.	Area	Co.	C	L of P
-----------------	----------	-------	------	----------	-----------	------	-----	---	--------

LOCATION OF ESTABLISHMENTS

(No. of estab.) (City) (County) (State)

(ABOVE IS MAILING ADDRESS—CHANGE IF INCORRECT)

I. PRINCIPAL PRODUCTS MANUFACTURED OR KIND OF WORK DONE IN ORDER OF IMPORTANCE

II. EMPLOYMENT, PAY ROLL, AND HOURS. (Before entering data see Instructions on other side.)

YEAR AND MONTH	Period Covered ONE PAY PERIOD ONLY (preferably 1 week) ENDING NEAREST 15th OF MONTH (both dates inclusive)		WAGE EARNERS (see Instruction 1)				All Employees (see Instruction 2)		DO NOT USE Expl. Code (10)
	From— (2)	To— (3)	Number		Pay Roll (6) <small>AMOUNT INCLUDING OVERTIME, VACATION, AND HOLIDAY PAY (see Instruction 3) (omit cents)</small>	Hours (7) <small>TOTAL NUMBER OF HOURS ACTUALLY WORKED, PLUS ESTIMATED PAID VACATION AND PAID HOLIDAY HOURS (omit fractions)</small>	INCLUDE ALL PERSONS WORKING ANY PART OF PERIOD, REGARDLESS OF TYPE OF WORK PERFORMED (include those in columns 4 and 5)		
			Both Sexes (4)	Women (5)			Both Sexes (8)	Women (9)	
(1)									
1943									
Dec.									
1944									
Jan.									
Feb.									
Mar.									
Apr.									
May									
June									
July									
Aug.									
Sept.									
Oct.									
Nov.									
Dec.									

III. OPERATING DAYS, VACATIONS, WAGE-RATE CHANGES AND COMMENTS.

YEAR AND MONTH	OPERATING DAYS		VACATIONS PAID FOR BUT NOT TAKEN		GENERAL WAGE-RATE CHANGES AND INDIVIDUAL CHANGES FOR LENGTH OF SERVICE, MERIT, OR PROMOTION			COMMENTS <small>Indicate below the main factors responsible for changes in employment, pay roll or hours from preceding month, such as: Seasonal expansion or contraction, inventory, repair, strike, holiday, material shortage, more or less overtime, increase or decrease of piece work, shift in type of labor, etc.</small>
	NUMBER OF DAYS WORKED BY MAJORITY OF EMPLOYEES DURING:		Number of Wage Earners (14)	Amount of Payments (15)	IF ANY SINCE LAST MONTH'S REPORT PLEASE CHECK		NUMBER OF WAGE EARNERS AFFECTED (18)	
	Week Ending Nearest the 15th (12)	Pay-roll Period Reported Above, if Different From Col. 12 (13)			Increase (16)	Decrease (17)		
(11)								(19)
1944								
Jan.								
Feb.								
Mar.								
Apr.								
May								
June								
July								
Aug.								
Sept.								
Oct.								
Nov.								
Dec.								

(OVER)

(Signature of person making report)

(Position)

16-37587-1 U. S. GOVERNMENT PRINTING OFFICE

From the CES inception up until 1983, mail served as the primary data collection instrument. A paper form, called the “mail shuttle,” was printed and mailed to each sample unit at the beginning of the year, with space for all 12 months of employment, earnings, and hours data. (See figure 1.) After entering data for a given month, respondents mailed the form to the appropriate BLS-funded state workforce agency, where it would be read and tabulated before being returned to the respondent for entry of the following month’s data. In this environment, collection was extremely disaggregated, because nearly identical functions were being carried out in all states and the District of Columbia.

Given the short collection period of the survey, the mail shuttle was typically able to collect only 40–50 percent of the expected reports for first preliminary estimates.⁶ By the time final estimates were published (2 months after the preliminary release), response rates would eventually approach 90 percent. Mail collection was particularly problematic for data that were flagged for edit reconciliation.⁷ States had to review such data in order to determine the source of the error and, in some cases, had to follow up with respondents. Given the limited time for receiving and processing information, these additional steps would almost certainly prevent data from being included in the initial preliminary estimates.

The mail-shuttle collection method was costly, because forms had to be mailed twice each month. Further, the forms were processed manually, which required transcription, tabulation, edit reconciliation, and followup with nonrespondents. Despite the shortcomings of the mail shuttle, it was the best collection instrument for a number of decades, given the technology of the period.

As the U.S. economy grew rapidly in the years following the end of the Great Depression and World War II, employment statistics received more attention. With the growth in industries covered by CES, states in 1946 began printing on color-coded forms, to allow for rapid sorting.

Figure 2. CES data collection form, 1960

BLS Codes

Form BLS 790 B

Before entering data please see explanations on other side

BLS Codes			
State	Report No.	Ind.	

YEAR AND MONTHS	PERIOD REPORTED		DO NOT USE		EMPLOYMENT IN STATE OF				YOUR COMMENTS			
	PAY PERIOD One pay period only (preferably 1 week ending nearest 15th of month)	NUMBER OF DAYS during the pay period shown in cols. 2 and 3 on which the majority of employees worked plus holidays and vacation days used for (Nearst 15 day)	NUMBER		ALL EMPLOYEES		CONSTRUCTION WORKERS		Please enter in column 13 the main factors responsible for significant month-to-month changes in employment, average hours worked (col. 11 + col. 9), average hourly earnings (col. 10 + col. 11), etc., as indicated by this report. Examples are:			
			From— (Both dates inclusive)	Through (Both dates inclusive)	7 (1)	L/P	DO NOT USE	DO NOT USE	Enter in these columns the number of construction workers who worked during or received pay for any part of the period reported, the pay earned (before deductions), and all man-hours worked or paid for. Include pay and hours for overtime, sick leave, holidays and vacations	NUMBER OF CONSTRUCTION WORKERS (If none enter 0)	CONSTRUCTION-WORKER PAYROLL (Omit cents)	CONSTRUCTION-WORKER MAN-HOURS (Omit fractions)
1959	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Dec.									\$			
1960												
Jan.												
Feb.												
Mar.												
Apr.												
May.												
June.												
July.												
Aug.												
Sept.												
Oct.												
Nov.												
Dec.												

LOCATION OF WORK DURING 1960					
STATE	COUNTY OR SIMILAR AREA	DATE WORK BEGAN IN COUNTY	DATE WORK ENDED IN COUNTY	STATE	COUNTY OR SIMILAR AREA

(Person to be addressed if questions arise regarding this report)

(Position)

U.S. GOVERNMENT PRINTING OFFICE: 1959-O-518100

The basic process of CES data collection was largely unchanged over the next several decades. There were some minor changes to the format of the form and the data requested. The page orientation of the 1960 form was shifted to landscape, and this format persisted until 1996. (See figure 2.) The 1968 form dropped the collection of data for female production workers, a practice that continues to this day.⁸

Toward multimodal CES data collection (1984–present)

In the early 1980s, CES began researching methods that could improve survey response rates and data quality. Even the most efficient mail-shuttle collection process could gather only about half of the data in time for the production of first estimates of employment and wages. This limitation led to large revisions in subsequent months, as additional data were collected.⁹

In 1984, in its first evaluation of new collection methods, CES enrolled respondents as part of a personal visit and then randomly assigned them to report by either mail or phone. The initial series of phone interviews was done without computer assistance and required the manual tracking and entry of data.

Figure 3. CES CATI establishment enrollment screen

TOPCATI: Enrollment & Data Collection System, Version 1.1.1.5 - CES Survey, Collection Month: 11/2007

File Print Tools Help

Data Collection - Call Scheduled for Dec 05, 2007 06:00 AM

Save Close Notes Dial Scheduler Send Document

Enrollment Data Collection

Case ID: 060192350 Don't Print Data

NRP Code: 1 ENR pk#: 0

Pending Package: TDE pk#: 0

Preferred Delivery: WEB pk#: 1

Advance Notice: EML pk#: 1

Dcc Code: 71 Refusal Panel #: 1

Refusal Conversion: 1

Primary Contact Secondary Contact

DC	Unit ID	Report With	Agg	Panel#	NAICS	Own	Form	Size	Av Ann AE	UI Number	RUN	EI Number	MSA	County	City Cd	Samp Code	Si
99	060192350	060192350		0073	238911	5	B	3	1234567890	00000	123456789	00000	995	000	2		

00 << Assign DC to All Units Enrollment Complete Zoom Out 1 of 1 units

Unit ID: 060192350 Physical Address

Company: Small Construction Company

Sec Name: Small Construction Company

Location: 12 Main St

City, State: Carmel CA

Zip Code: 99888 Type: CA

Panel Number, Sub Sample: 0073

County, City, MSA, CMI: 995 000 16

NAICS, Own, Size, Form: 238911 5 3 B

EI Number, Aux, MEEI: 123456789 5 1

UI Number, RUN: 1234567890 00000

Sample Code, Unit Code: 2 AL

Pr Lopp, Cm Lopp: 1 0

QUI Data

Month/Year	AE	AE IMP	Wage	Wage IMP	Comment
03/2002	19	R	101462	R	
02/2002	19	R			
01/2002	19	R			
12/2001	18	R	109300	R	
11/2001	18	R			
10/2001	18	R			
09/2001					
08/2001					

Unit Addresses Unit Info

Ready

Figure 4. CES CATI monthly data collection screen

TOPCATI: Enrollment & Data Collection System, Version 1.1.1.5 - CES Survey. Collection Month: 11/2007

Data Collection - Call Scheduled for Dec 05, 2007 06:00 AM

Save Close Notes Dial Scheduler Send Document

Enrollment Data Collection

Select 1 of 1 units RC

060192350 Small Construction Company 00 Case ID 060192350

Contact Bob Small Location State CA NAICS 238911 Form Type B Mode W-R

CES Microdata

AE/PW/MM/YYYY	Empl	WW	PR	Comm	HR	OT	Plp	Clp	CC1	CC2	GME	GCC	RC
AE 11/2007							1	0					00
PW													
AE 10/2007	10	2	5000		400		1	0					90
PW	8		4000		320								

AE/PW/MM/YYYY	AWH	AHE	AWE	AOT	PW/AE	WW/AE	GME/AE prev per wk	Good Date
AE 11/2007								
PW								
AE 10/2007	40.00	12.50	500.00		0.8000	0.2000		11/06/2007
PW	40.00	12.50	500.00					

Open Multiple Payroll How many persons worked or received pay during the pay period of November 12th? Roll Over

Start Interview

Select Option to begin Interview

☒ Dial ☐ Don't Dial

TDE Phone: 800-877-7715
Help Line: 800-827-2005
Number of Missing Months: 0
Number of Calls Placed: 0
Last Call:

Cancel << Back Next >>

Ready

Computer-assisted telephone interviewing (1985). After initial testing success, BLS began evaluating the suitability of computer-assisted telephone interviewing (CATI) for CES data collection. CATI software is designed to provide prompting information and data capture to survey interviewers collecting information by phone. (See figures 3 and 4.) The computer program maintains call lists, allows for autodialing, and then provides interviewers with prompts that guide data collection. If the collected data do not meet expectations during data entry into the system, the computer prompts interviewers to follow up with more questions. Additional visual cues can also be provided to interviewers, to allow for the live editing of data.

In 1985, BLS used CATI software developed by the University of California at Berkeley in an initial test targeting 370 CES respondents in Maine and Florida. The test sample was representative of the CES survey for these states and included a range of industries and business establishments of various sizes. During the test, more than 95 percent of respondents agreed to provide data by phone on a monthly basis. Even more impressive was that the response rate at first preliminary estimates (i.e., data that would be included in the initial national release of employment estimates) went from 45 percent to 85 percent. Further, the CATI method required only an average of 1.5 calls, totaling 5 minutes, for successful data collection. Initial testing aimed at determining and setting collection procedures and was performed by states under BLS oversight.

The next phase of testing was expanded to include nine states. Of those, six states (Alaska, Iowa, Maine, Missouri, Nebraska, and Vermont) tested a representative subsample and three states (Alabama, California, and Florida) targeted respondents that were consistently late in returning their mail-shuttle forms. These tests

included 3,300 sample units in 1987 and, after the addition of two more states in 1988 (Georgia and Mississippi), an additional 2,500 sample units. The goal of this round of tests was to measure interviewer productivity, staff utilization, and cost. This information would help determine the practicality of a full implementation of CATI data collection.

The expanded series of tests enjoyed much of the same success as the initial round, with first-month collection rates at 85 percent (nearly double the mail response rates). Interviewers needed only an average of 1.7 calls to collect data, spending a total of 4.4 minutes on each case (2.5 minutes of which was on the phone). Perhaps not surprisingly, attrition was cut in half. While respondents could tacitly refuse to complete and return their forms under the mail-shuttle system, they had to overtly refuse to provide data during a phone-based CATI collection.¹⁰

The CATI tests highlighted the advantages of phone collection over mail collection and resulted in a paradigm shift for CES. The new model of actively collecting data by phone would fully harness new technologies available to both BLS and respondents.

Collecting CES data by CATI, while relatively costly, is advantageous for a number of reasons, none more so than the timeliness of collection. CES respondents cannot submit data before their firm's reference pay period has been processed. Business payrolls can be weekly, biweekly, semimonthly, or even monthly. This variation is a major challenge for a quick-response survey, because respondents may have only a limited number of days from the time payroll is processed until CES completes its data collection for first preliminary estimates. In fact, depending on the calendar, monthly payrolls are frequently not available before this early collection deadline. (Recent research by CES has determined that 11.3 percent of the CES sample for private businesses is collected from monthly payrolls.¹¹) This time constraint makes mail collection problematic. The majority of respondents simply do not have enough time to process payroll, complete mail forms, and return these forms in time for first preliminary estimates. By contrast, CATI collection excels as a quick-response data collection instrument, because phone calls can be scheduled to meet each firm's payroll processing schedule. Not only are respondents considerably more likely to provide information consistently, but they are also much more likely to provide it before CES runs its first monthly estimate.

In addition to being successful in collecting data at high rates (see appendix), CATI collection allows for the instantaneous edit and screening of data. Sophisticated CATI systems can flag data that either fail a logical edit (e.g., presence of more female employees than total employees) or a screening check (e.g., presence of high average earnings for a particular industry). As the interviewer enters data collected over the phone, the computer system immediately flags suspect data for followup questions. Flagging minimizes respondent burden (by eliminating, or at least reducing, followup inquiries) and ensures the collection of accurate data. Results from feasibility tests indicate that more than half of cases are completed during a single phone call and fewer than 20 percent of cases require more than two phone calls (these figures include calls not answered and busy signals).

In 1990, after successful CATI testing, BLS opened its first CES data collection center (DCC) in the Atlanta regional office. Initially, this DCC was modestly staffed with three contract interviewers and a first-line supervisor overseen by a federal manager. DCC operations grew over time as similar centers were opened in Kansas City in 1992, Dallas in 1997, and Niceville, FL, in 2004 (the Niceville DCC has since moved to Fort Walton Beach, FL).¹² Staff levels increased considerably as CATI became the primary method for enrolling new respondents

and collecting data. There are now more than 300 contract staff working at these DCCs. Some staff work only part time during the second half of the month, for peak collection. In recent years, DCCs have added bilingual interviewers to allow for the enrollment of Spanish-speaking respondents.

DCCs are responsible for the following steps in data collection:

- (1) *Address refinement.* After receiving a monthly panel of sample units, DCCs are responsible for confirming the location of each unit and identifying a point of contact for CES survey enrollment.
- (2) *Survey enrollment.* A DCC interviewer sends an information packet to the identified business establishment contact, providing an overview of the CES program and a sample form. Shortly after this packet is received, the interviewer calls the potential respondent and attempts to enroll his or her establishment in the CES survey.
- (3) *Data collection.* After enrollment, the interviewer becomes responsible for monthly data collection. The interviewer typically schedules a CATI call for a time when payroll data are available. During this call, the interviewer collects data and, if necessary, asks followup questions to ensure that the data pass edit and screening tests or have the appropriate comment codes assigned. These comment codes are later used by estimation staff in the review of data.
- (4) *Collection rolling.* Because of the relatively high cost of CATI data collection, CES prefers to transfer establishments to a self-reporting method after 5 months of CATI collection. The interviewer is responsible for identifying respondents who would be a good fit for self-reporting and then for providing guidance on the transfer. Some establishments, however, remain in a permanent CATI status if the interviewer considers them ineligible for self-reporting or if they request such status.
- (5) *Edit reconciliation.* For data that are self-reported and fail edit and screening tests, a group of interviewers is responsible for performing edit reconciliation. This process involves a followup with the respondent to determine if the data are correct and, if so, to properly document the reasons for the unexpected data change.

DCCs produce strong returns, but CATI operations entail considerable labor costs. CATI collection has proven to be very effective in enrolling new respondents and “training” them in providing CES data, but its high cost constrains the sample size CES can maintain. In order to maximize the returns of our nation’s “data dollar,” CES had to develop a more efficient method to collect data from respondents who had been initially enrolled in CATI. It was clear that respondents were becoming knowledgeable about reporting data during their several months of working with a CES data collector. This fact suggested that an eventual shift to a self-reporting method could lower costs without sacrificing data quality.

Touchtone data entry (1987). In 1987, the CES program began testing touchtone data entry (TDE) as a collection instrument.¹³ This system provides respondents with a toll-free number that is available around the clock. The caller enters an identification number (known in CES as a “report number”), triggering an automated interview. The respondent is given a computer-based interview that replicates the entries that appear on paper forms or the standard questions asked by interviewers during a CATI call. The respondent then provides the

data by selecting the correct keys on his or her touchtone telephone. For verification purposes, the answers are read aloud by the computer. When the call is complete, the data are submitted to CES.

After successful testing for a small number of respondents in Maine and Florida, TDE was implemented nationwide. Each state maintained a toll-free number and managed its own help-desk and edit-reconciliation processes. To streamline data collection, CES did not perform editing during the initial automated phone call. TDE was used only for respondents who had been “trained” during 5 months of CATI collection. If, at 6 months, the CATI interviewer determined that a respondent was sufficiently reliable and knowledgeable about the reporting process, that respondent would be transitioned to TDE reporting. If a respondent was either a poor fit for TDE or unwilling to adopt the new method, then collection would continue by CATI.

In 2004, the toll-free number, as well as the help-desk and edit-reconciliation processes, moved away from individual states and became centralized at BLS, where everything remains today. This transition resulted in substantial efficiency gains and cost savings for the program. Unlike CATI collection, in which BLS interviewers provide edit and screening during a CATI phone call, TDE relies on a followup phone call only if a TDE response fails edit and screening tests. In 2012, CES added a Spanish-language option for TDE.

Under the TDE system, respondents receive a monthly fax or postcard reminder to report their data. This practice has proven to increase collection rates. TDE collection rates have risen to 85 percent in recent years (see appendix), but when the method was used on a larger scale (with less selectivity of respondents), collection rates were typically between 70 and 75 percent. These rates are lower than those for CATI, but they come at a substantially lower cost (TDE has less than half the per-unit cost of CATI). At peak usage in the 1990s, more than 30 percent of the CES sample was collected by TDE. In 2005, CES began requesting additional data items, which made TDE too cumbersome for most respondents. Currently, TDE is used to complete only 3 percent of CES reports and is reserved mostly for state and local government respondents (these respondents have fewer data items to complete) and for special requests.

Fax (1995). With fax machines becoming ubiquitous in business during the early 1990s, CES in 1995 introduced a faxable version of the traditional form for respondents with multiple worksites. Each month, CES faxed a blank form to respondents, asking them to complete the form and return it by fax. Upon fax receipt, DCC staff would keypunch the data into the CES CATI system. Fax collection proved to be a reliable method, registering an 85-percent collection rate by first preliminary estimates. (See appendix.) This rate was comparable to that of CATI, but achieved at a lower cost. The existing fax process still requires human effort for data entry. However, the process remains relatively efficient, because respondents can provide data for a large number of worksites in a single response.

Electronic data interchange (1995). As it became more common for firms with multiple worksites to maintain large central computer databases with payroll information, BLS developed a procedure for using these data. In February 1995, it opened an electronic data interchange (EDI) center in Chicago, with the purpose of transforming the large datasets into usable data for CES and another BLS federal–state program, the Quarterly Census of Employment and Wages (QCEW).

The EDI method has large upfront costs, because it requires a team of economists, analysts, and programmers to process and analyze massive data files provided voluntarily by firms. Using proprietary software developed

by BLS, analysts transform these firm files into data usable by CES and QCEW. The large upfront costs associated with this effort are balanced by low monthly processing costs. On a monthly basis, the EDI staff ensures that the data files being processed conform to the established format and then performs edit reconciliation on worksites that are flagged for potential errors.

The EDI method provides tremendous value—more than 40 percent of CES worksites are processed with relatively limited respondent burden and low costs for BLS. (See appendix.) However, this method depends heavily on data being available to BLS before the deadline for first preliminary estimates. Collection rates for EDI respondents could be highly volatile, particularly in short months, because a lack of data from a few major firms can affect overall CES response (as data become available, these firms are included in later revisions of the Employment Situation). Further, only a small percentage of firms have enough worksites to make EDI collection an efficient option. Previous research by CES has indicated that EDI is a practical option for only about 5 percent of U.S. firms.

Figure 5. CES Web collection

The screenshot shows the CES Web collection interface. At the top is the Bureau of Labor Statistics logo and navigation links for Help and Logout. Below this is a section for the CES Report Number and Pay Group 1. The main form is titled 'Report the columns below for the two week payroll that includes the 12th of' and contains a table with five columns: Employee Count, Women Workers, Payroll, Excluding Commissions (Whole dollars), Commissions Paid at Least Once a Month (Whole dollars), and Hours, Including Overtime (Whole hours). The form also includes a 'Pay' dropdown set to 'Bi-weekly' and a 'Commissions' dropdown set to 'No Commissions'. Below the table are links for 'All Workers', 'Production Workers', and 'Reason for Large Changes'. At the bottom is a 'Continue' button.

	1	2	3	4	5
	Employee Count	Women Workers	Payroll, Excluding Commissions (Whole dollars)	Commissions Paid at Least Once a Month (Whole dollars)	Hours, Including Overtime (Whole hours)
All Workers					
Production Workers					
Reason for Large Changes	Use for Columns 1-5 (payroll, hours and employment) <div> none none </div>				

Continue →

Web-based collection (1996). As computer usage and Internet availability expanded in the 1990s, CES became the first federal program to test Web reporting. After a small-scale test in 1996, CES decided in 1998 to begin full production of a Web-based data collection application. (See figure 5.) For respondents who had access to a computer and the Internet, the Web application would provide an alternative to TDE, which had been the predominant self-reporting method for the previous decade.

Web respondents are prompted by email when it is time for them to provide data. If data are not submitted, reminder emails are sent as a followup. Email prompting proved more advantageous than the mail and fax reminders of TDE, because its low cost and low perceived burden allowed for more contact with respondents. CES also found that, when presented with a Web interface displaying a visual cue (similar to paper forms) of the full request, Web respondents were more likely to provide all requested data items. Compared with other self-reporting methods, Web collection had the advantage of being able to display edit-reconciliation prompts to the respondent during data entry. If an entry failed edit or screening tests, the respondent would be prompted to

review the entry and either correct the reported data or provide a comment or an explanation code. This process substantially reduced the number of followup contacts with the respondent, minimizing respondent burden and providing more usable data for preliminary estimates.

Despite its advantages, the Web method consistently trailed TDE in collection rates. One major issue identified by CES was that respondents had trouble maintaining login passwords. The issue was particularly problematic in cases with respondent change, a frequent occurrence when firms remain in sample for a period of 2 to 4 years. In 2006, CES tested a new Web design, called “Web-lite,” that did not require a password. Under the new design, respondents access the Web-lite page by clicking on a link (sent by CES in an email) and completing a CAPTCHA.¹⁴ Because the application is no longer password protected, firm-identifying information or data entries that were required in earlier designs are not requested. This feature limits edit and screening requirements at this stage of the process. Full edit and screening tests are performed internally by BLS after Web data are imported. Entries that fail these tests are sent through an edit-reconciliation process that involves a followup phone contact with the respondent for confirmation and explanation of changes. Further, respondents cannot see their contact information displayed, so it is not always clear when this information needs to be updated. The system is designed to provide semiannual reminders to respondents, asking them to update their contact information. Despite these issues, Web-lite tests indicated far better collection rates and lower overall respondent burden than earlier Web-based methods. After completing the tests, CES adopted Web-lite as the standard Web-entry application.¹⁵ A Spanish-language version of Web-lite was created in 2011.

Web collection rates are typically around 80 percent for first preliminary estimates, trailing CATI collection rates by only about 10 percentage points. (See appendix.) However, data from Web respondents can be collected at about a third of the cost of CATI collection—a considerable cost saving. The success of Web is largely due to the initial CATI collection period. During this period, respondents are screened (to ensure that Web is the appropriate data collection instrument for their situation) and trained in providing data and establishing a monthly routine. Similarly to TDE, Web collection is considered a complement to, not a replacement for, CATI collection and enrollment.

Modernizing the paper form (1996). Seeking to catch up with the times, in 1996 BLS modernized the mail form with a sans serif typeface and space for collecting an email address. After decades of asking respondents to provide the start and end dates of their pay periods, the mail form introduced checkboxes for selecting the length of the pay period. Another major addition was a comment-code section. While previous versions of the form had allowed respondents to provide open-ended comments, the modernized version was the first to provide a list of comment codes for selection. This comment-code redesign permitted much faster processing. Although mail was largely replaced by modern data collection instruments in the late 1990s, CES continues to use paper forms in mailed enrollment packages and for replacement forms (when the enrollment forms have been completed). Paper forms are no longer used as a mail shuttle, but rather as a method for allowing respondents to view instructions, visualize data requests, and maintain records of their responses. The forms also include BLS contact information, providing respondents with a single point of contact for any questions.

In 2008, CES made further improvements to the form.¹⁶ After research and testing, the program adopted a redesign consisting of a single 11" × 17" piece of paper folded to form four pages. The first page is a cover letter providing BLS contact information and informing a business establishment that it has been selected by the CES

Figure 6. Redesigned CES enrollment form

Your Report Number is: Pay Group 1

► When you report your payroll for the pay period that includes the date of **January 12th**, we will only ask for answers to these questions.

► Each following month, we will call for the pay period that includes the 12th of each month. You can keep a record of what's reported each month below:

We will send you another form for reporting after June 2015.
Please keep this form to use when the Data Collection Specialist calls you to complete the survey. Thank You!

13

changes. The most substantial form modification is the visual separation of the first month of reporting from subsequent months. (See figure 6.) This modification lessens the visual burden on the respondent and signals that only 1 month of data needs to be completed at a time.

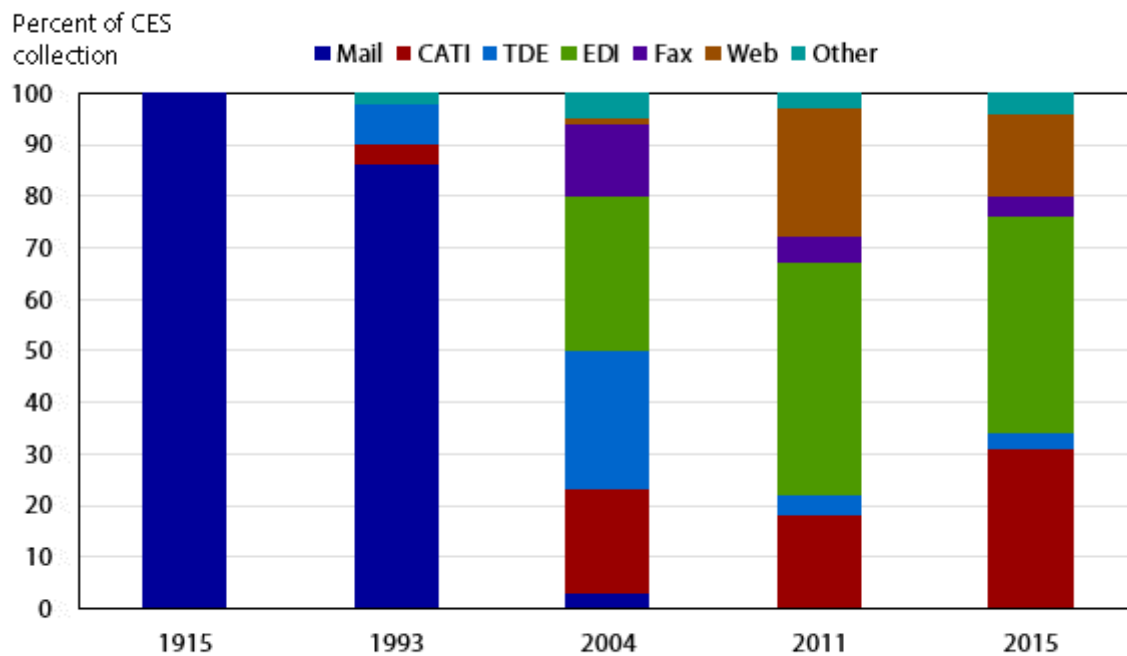
During testing, CES compared the response rates achieved with the new form with those of the previous forms. Although little impact was observed for unit response rates, the new form proved to increase item response rates.¹⁷ The booklet print format also eliminated the possibility of forms from different firms being mismatched during assembly and mailout.

Collection by spreadsheet (2007). In recent years, as the use of fax by businesses has declined, CES has introduced alternative methods to collect data from midsize firms. One process, called WebFTP, involves the completion of a preformatted Excel spreadsheet. Upon completing the spreadsheet, respondents upload it to the CES website with a file transfer protocol. This process, which CES began using in 2007, was originally developed by West Virginia, before the bulk of data collection responsibility shifted to BLS. Currently, the method is the collection instrument with the lowest respondent burden for midsize firms. However, despite its promise as a low-cost instrument, WebFTP has been problematic for BLS because the processing of spreadsheets has proven rather time consuming in practice. The spreadsheets' flexibility has made their automated reading difficult, as respondents frequently make changes to the data format. In addition, respondents often provide notes with the expectation that their responses are reviewed by a human. Currently, CES is working to develop improved procedures for this process and, ideally, develop a replacement data collection instrument for midsize firms.

A century of improving CES data collection

The CES survey has evolved substantially over the past 100 years. Initially, it used a mail-shuttle form to cover four industries with several hundred sample units. Today, it uses diverse collection methods and covers all nonfarm industries, representing more than 500,000 sample units. This modern multimodal survey maximizes response by reducing respondent burden while ensuring the highest quality of data. As a quick-response survey, it collects as much data as possible in time for first preliminary estimates. These estimates are part of the Employment Situation, typically released on the first Friday of each month. Although collection for first preliminary estimates ends around the turn of each month, CES continues to gather additional data for revisions that occur over the next 2 months.

Figure 7. CES data collection methods, 1915–2015

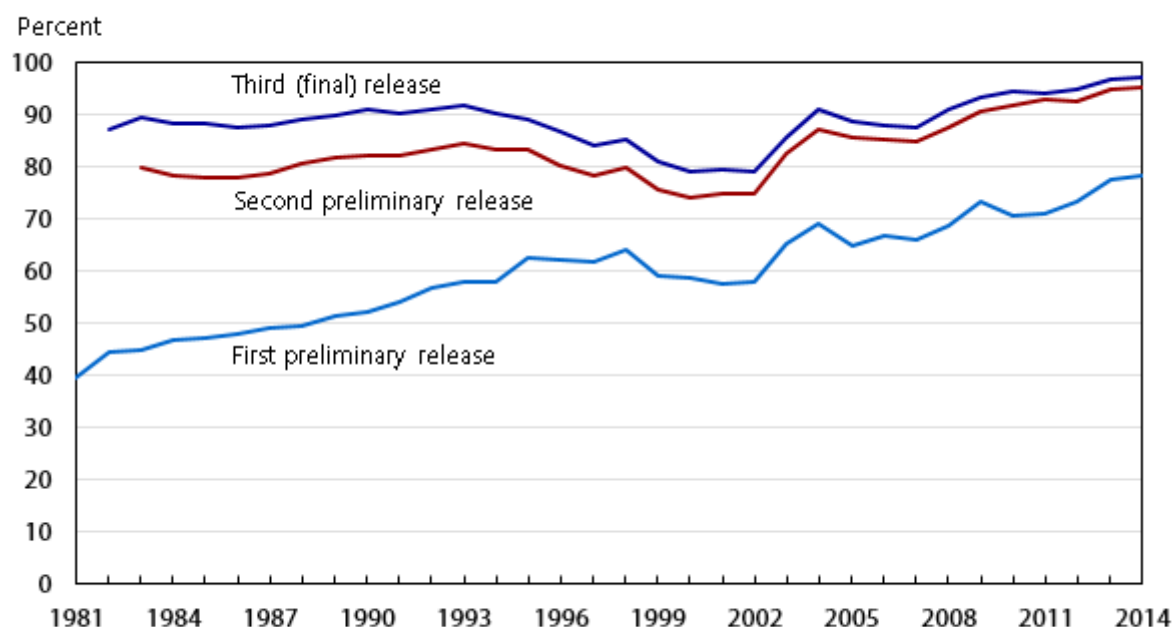


Note: The years shown represent major inflection points in the transition over time from mail to multimodal CES collection.

Source: U.S. Bureau of Labor Statistics.

The evolution of CES data collection over the past 100 years clearly demonstrates the shift in methodology.¹⁸ (See figure 7.) As the CATI and TDE methods entered production in the late 1980s, they started chipping away at the mail shuttle. By 2004, CES efforts to offer innovative collection instruments clearly began to show, as CATI, TDE, EDI, Web, and fax almost completely displaced mail collection. By 2011, Web had taken over most of TDE collection, and EDI continued to expand. An interesting change, caused by a shift in program resources, came about in 2015, when CATI increased as a percentage of total CES data collection.¹⁹ Expansion occurred in both the Kansas City and Fort Walton Beach DCCs, to allow for the collection of data from a greater number of permanent CATI respondents. (A permanent CATI respondent is simply a respondent who continues CATI reporting after 5 months of initial collection.) This extended collection could occur for various reasons, including the size of the establishment, the respondent's willingness to report through CATI, or the complexity of data that need to be collected. The CATI expansion was based on CES research indicating that a number of establishments that were being rolled over to Web were experiencing considerable dips in collection rates. CES identified certain firm-size classes and industries that responded better to CATI collection than to Web and provided the additional CATI resources to allow for this collection.

Figure 8. CES collection rates, 1981–2014



Source: U.S. Bureau of Labor Statistics.

After nearly 70 years of relying on a single mail-shuttle form, the CES program devoted substantial resources to modernizing and enhancing the data collection process. In particular, the program sought to improve collection rates for first preliminary estimates, the earliest read on employment. Figure 8 clearly demonstrates the gains that have been made over the past 30 years with the addition of the CATI, TDE, EDI, fax, and Web instruments. Collection rates for the national first preliminary estimates have nearly doubled, reaching almost 80 percent by 2015 and substantially reducing later data revisions. Gains have also been made for the second preliminary release, which occurs 1 month after the first preliminary release. In the last three decades, collection rates for the second release have risen from approximately 80 percent to approximately 95 percent. Finally, by the time CES reaches its third, and final, release (2 months after the first preliminary release), approximately 97 percent of data have been collected.

The future of CES data collection

To enhance data collection methods, CES experts are constantly working to evaluate current processes in light of new technologies. Several efforts in this direction are underway.

Collection of data from mid-sized firms with multiple worksites remains the most challenging effort for CES. Currently, research is in progress to determine if an updated Web collection instrument, which would target respondents that report for five or more worksites, would be helpful in reducing respondent burden. While the current Web form is appropriate for respondents with fewer than five worksites, it is inadequate for respondents with five or more worksites. The new Web form, which is in the early stages of development, would allow for the entry of more than one worksite on a single page, reducing the burden on multiworksite respondents.

Other ongoing efforts aim at reducing data collection risks. The quick-response nature of the CES survey, although partially ameliorated by multimodal diversification, is still susceptible to the inherent risks posed by extreme weather and equipment failure. Power outages, hurricanes, and winter weather—events that can close a DCC or limit its operations—are just a sampling of the types of problems that can occur during peak CES collection. A current project is seeking to develop a fax-based optical character recognition (OCR) system that would allow machine reading of incoming faxes. The new system would enable CES to collect data even during a DCC closure. As one illustration, imagine that a hurricane closed a DCC for the final week of data collection. With the click of a button, a broadcast fax system could send faxes to all respondents from the impacted DCC, requesting that data be provided by fax. As the faxes were returned, the BLS computer system would employ modern OCR technology to read and accept the data directly into the CES processing system. CES software would then run edit and screening tests to determine which data should be included in estimation.

One collection method that has proven somewhat elusive in recent years has been email data collection. There has been demand from respondents, particularly over the past 15 years, to report CES data by email. Although some informal collection by email exists, there has been no formally established collection procedure for email. In 2007 and 2008, CES performed research to design an email collection instrument that would allow respondents to provide data in a form located in the body of an email. Unfortunately, testing of this process identified a number of insurmountable problems, chief among which was the lack of consistent reporting among email clients. A new method relying on a fillable PDF form is now in development and may be deployed in the near future. The method could be used both for regular data collection and in conjunction with the emergency fax solution detailed earlier.

The CES program continues to evaluate current data collection instruments. The transition to a multimodal survey over the past 30 years has considerably improved collection rates for preliminary estimates. As a result, data revisions have been reduced substantially. Collectively, these efforts and improvements allow for a clearer picture of the state of our economy, at the earliest possible time—a clear benefit for our entire nation. As the CES program celebrates a milestone birthday, we would like to thank our respondents for providing data over the past 100 years. These data allow BLS to provide key decisionmakers, and the nation, with the most accurate gauge of the labor market.

Appendix

Collection costs and rates, by collection method

Collection method	Cost	Typical first release collection (percent)	Typical final release collection (percent)	Current percent of CES collection	Period of peak CES collection
Computer-assisted telephone interviewing (CATI)	\$\$\$ \$\$	90	97	31	2010–15
Touchtone data entry (TDE)	\$\$	85	97	3	1995–2005
Fax	\$\$\$ \$	85	97	4	1995–2006
Electronic data interchange (EDI)	\$	60	95	42	2010–15
Web-based collection	\$\$	80	93	16	2013

See footnotes at end of table.

Source: U.S. Bureau of Labor Statistics.

NOTES

¹ For more information, see <http://www.bls.gov/ces/>.

² The Employment Situation presents statistics from two major surveys, the Current Population Survey (household survey) and the Current Employment Statistics survey (establishment survey). The household survey—a sample survey of about 60,000 eligible households conducted by the U.S. Census Bureau for BLS—provides information on the labor force, employment, and unemployment. The establishment survey provides information on employment, hours, and earnings of employees on nonfarm payrolls. BLS collects these data each month from the payroll records of a sample of nonagricultural business establishments.

³ In 1915, the CES sample covered fewer than 500 business establishments. Today, it includes 143,000 businesses and government agencies, representing 588,000 unique worksites.

⁴ The CES survey is typically considered a quick-response survey because of the short turnaround time between the program's reference period and data release.

⁵ CES data are also used to produce state and area estimates, which are typically released 2 weeks after the release of national estimates. These state and area estimates are revised in the following month before being considered final. This revision differs from that for national estimates, which are revised twice, both in the first month and the second month after initial release.

⁶ For historical collection rates, see <http://www.bls.gov/web/empsit/cesregrec.htm>.

⁷ CES performs checks on collected data to identify potential errors. Data flagged during these checks are verified with the respondent during edit reconciliation. Data are not used in estimates until verification has occurred.

⁸ For more information on the history of CES forms, see Louis J. Harrell, Jr., and Edward Park, "Revisiting the survey form: the effects of redesigning the Current Employment Statistics survey's iconic 1-page form with a booklet style form," *Proceedings of the Joint Statistical Meetings* (American Statistical Association, October 2012), pp. 4900–4914, <http://www.bls.gov/osmr/pdf/st120260.pdf>.

⁹ For more information on CES revisions, see Thomas Nardone, Kenneth Robertson, and Julie Hatch Maxfield, "Why are there revisions to the jobs numbers?" *Beyond the Numbers*, vol. 2, no. 17, July 2013, <http://www.bls.gov/opub/btn/volume-2/revisions-to-jobs-numbers.htm>.

¹⁰ For more information on the initial testing of CATI collection for CES, see George Werking, Richard Clayton, Richard Rosen, and Debbie Winter, "Conversion from mail to CATI in the Current Employment Statistics survey," *Proceedings of the Survey Research Methods Section* (American Statistical Association, 1988), pp. 431–436, https://www.amstat.org/Sections/Srms/Proceedings/papers/1988_079.pdf.

¹¹ For more information on the pay frequency of private businesses in the CES sample, see Matt Burgess, "How frequently do private businesses pay workers?" *Beyond the Numbers*, vol. 3, no. 11, May 2014, <http://www.bls.gov/opub/btn/volume-3/how-frequently-do-private-businesses-pay-workers.htm>.

¹² For more information on the organization and management of a DCC, see Paul Calhoun, Laura Jackson, and John Wohlford, "Organization and management of a data collection center," *Proceedings of the Joint Statistical Meetings* (American Statistical Association, August 2011), <http://www.bls.gov/ore/pdf/st010050.pdf>.

¹³ Voice recognition (VR) software was also tested in 1989 and produced similar results to TDE. As testing was underway, the use of touchtone phones became near universal, so the need to update TDE with VR software was no longer considered cost effective. Therefore, voice recognition technology was never implemented.

¹⁴ CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a method for ensuring that it is a human and not a computer that is attempting access to an online application. CES has gradually shifted from using a text-based CAPTCHA to using a picture-based CAPTCHA, which respondents find more user-friendly.

¹⁵ For more information on CES Web collection, see Richard J. Rosen, Christopher D. Manning, and Louis J. Harrell, Jr., “Web-based data collection in the Current Employment Statistics survey,” *Proceedings of the Survey Research Methods Section* (American Statistical Association, 1998), pp. 354–359, https://www.amstat.org/sections/srms/Proceedings/papers/1998_058.pdf. For more information on BLS Web collection, see Stephen Cohen, Dee McCarthy, Richard Rosen, and William Wiatrowski, “Internet collection at the Bureau of Labor Statistics: an option to report data,” *Monthly Labor Review*, February 2006, pp. 47–57, <http://www.bls.gov/opub/mlr/2006/02/art4full.pdf>.

¹⁶ This work was conducted under contract with Dr. Don Dillman of the Social and Economic Sciences Research Center at Washington State University.

¹⁷ Unit response rate refers to the percentage of establishments (or units) that responded with at least partial data. Item response rate is calculated for each individual data item requested from establishments. For example, CES requests information on employee count, female-employee count, payroll, commissions, hours, and overtime for both all workers and production workers (female-employee count is requested only at the all-employee level). CES measures response rates for each of these data items because not all establishments are able or willing to provide data for each individual request.

¹⁸ For more information on the expansion of CES data collection over the past 100 years, see Richard L. Clayton, “Implementation of the Current Employment Statistics redesign: data collection,” *Proceedings of the Survey Research Methods Section* (American Statistical Association, 1997), pp. 295–297, http://www.amstat.org/sections/srms/proceedings/papers/1997_048.pdf; Kenneth W. Robertson and Julie Hatch Maxfield, “Data collection in the U.S. Bureau of Labor Statistics’ Current Employment Statistics survey” (Geneva, Switzerland: United Nations Economic Commission for Europe, Conference of European Statisticians, 2012), <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2012/mtg2/WP20.pdf>; and Richard J. Rosen and David O’Connell, “Developing an integrated system for mixed mode data collection in a large monthly establishment survey” (U.S. Bureau of Labor Statistics, 1997), <http://www.bls.gov/ore/pdf/st970150.pdf>.

¹⁹ For more information, see <http://www.bls.gov/sae/cesprocs.htm>.

RELATED CONTENT

Related Articles

[The Monthly Labor Review turns 100](#), *Monthly Labor Review*, July 2015.

[How frequently do private businesses pay workers?](#), *Beyond the Numbers*, May 2014.

[Why are there revisions to the jobs numbers?](#), *Beyond the Numbers*, July 2013.

Related Subjects

[History](#) | [Bureau of labor statistics](#) | [Survey methods](#) | [Technological change](#) | [Statistical programs and methods](#)